(19) JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **09152892 A**

(43) Date of publication of application: **10.06.97**

(51) Int. Cl **G10L 5/04**
G10L 3/02
G10L 7/02
G10L 9/00

(21) Application number: **08238235**

(22) Date of filing: **09.09.96**

(30) Priority: **26.09.95 JP 07248144**

(71) Applicant: **NIPPON TELEGR & TELEPH CORP <NTT>**

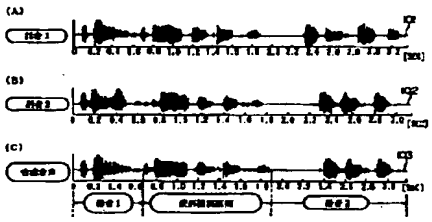
(72) Inventor: **ABE MASANOBU**

(54) **VOICE SIGNAL DEFORMATION CONNECTION METHOD**

(57) Abstract:

PROBLEM TO BE SOLVED: To provide a voice signal deformation connection method capable of connecting the voice messages of different voice quality each other without the sense of incongruity.

SOLUTION: Two voice signals 101 and 102 obtained by making two speakers utter the same text are connection-processed. Synthetic voice 103 resulted from this processing is constituted of the voice section of the speaker 1, a deformation connection section and the voice section of the speaker 2. Even in the case that the voice quality of the two speakers is widely different, when the extent of deformation at one time is small, a listener side does not feel the sense of incongruity so much. Then, the voice is connected by repeating the deformation without the sense of incongruity stopwise for several times. That is, in the voice message obtained by connection, the voice quality is gradually changed over the prescribed time of the deformation connection section.



(19)日本国特許庁（ＪＰ）　(12)公開特許公報（A）　(11)特許出願公開番号

特開平9－152892

(43)公開日　平成９年(1997)６月10日

(51)Int.Cl.*	識別記号	庁内整理番号	ＦＩ		技術表示箇所
G10L 5/04			G10L 5/04	D	
3/02			3/02	D	
7/02			7/02	D	
9/00			9/00	E	

審査請求　未請求　請求項の数12　ＯＬ　（全10頁）

(21)出願番号　特願平8－238235

(22)出願日　平成８年(1996)９月９日

(31)優先権主張番号　特願平7－248144

(32)優先日　平７(1995)９月26日

(33)優先権主張国　日本（ＪＰ）

(71)出願人　000004226
日本電信電話株式会社
東京都新宿区西新宿三丁目19番２号

(72)発明者　阿部　匡伸
東京都新宿区西新宿三丁目19番２号　日本電信電話株式会社内

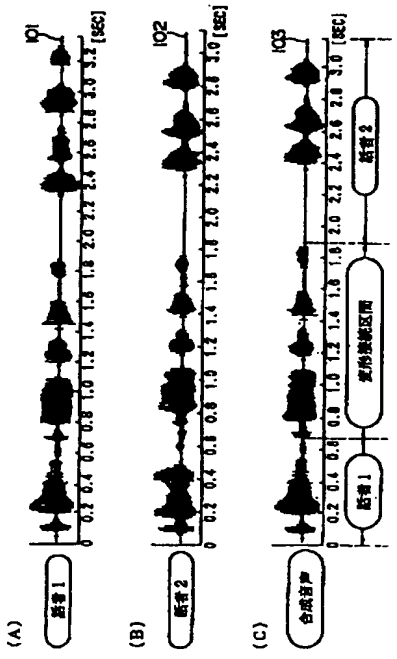
(74)代理人　弁理士　志賀　正武

(54)【発明の名称】音声信号変形接続方法

(57)【要約】

【課題】　声質の異なる音声メッセージ同士を違和感無く接続することができる音声信号変形接続方法を提供する。

【解決手段】　２人の話者に同じテキストを発声させて得られる２つの音声信号１０１及び１０２を接続処理する。この処理の結果生成された合成音声１０３は、話者１の音声区間、変形接続区間、話者２の音声区間とから構成される。２人の話者の声質が大きく違う場合であっても、一度に変形する度合いが小さければ、聞く側はあまり違和感を感じない。そこで、違和感を感じない程度の変形を段階的に複数回繰り返すことによって音声の接続を行う。すなわち、接続により得られる音声メッセージは、変形接続区間の所定時間に渡って徐々に声質が変化していく。



1

【特許請求の範囲】

【請求項１】　互いに異なる声質に属する２つの音声信号同士を接続する処理において、

音声信号の特性を示すパラメータの値を、一方の音声信号の特徴を示す値から他方の音声信号の特徴を示す値へと所定時間にわたって徐々に変化させることにより、２つの音声信号を接続することを特徴とする音声信号変形接続方法。

【請求項２】　請求項１記載の音声信号変形接続方法において、

前記パラメータの値を変更する所定時間にわたり、前記互いに異なる声質の話者に同一のテキストを読み上げさせ、これにより得られる２種類の音声データを用いて前記パラメータ値の変更を行うことを特徴とする音声信号変形接続方法。

【請求項３】　請求項１記載の音声信号変形接続方法において、

前記互いに異なる声質の音声信号は音声合成装置による発声により得られるものであることを特徴とする音声信号変形接続方法。

【請求項４】　請求項１記載の音声信号変形接続方法において、

前記互いに異なる声質の音声信号は、一方は人間による発声、もう一方は音声合成装置による発声により得られるものであることを特徴とする音声信号変形接続方法。

【請求項５】　請求項１記載の音声信号変形接続方法において、

前記パラメータは音声のスペクトルであり、前記所定の時間にわたって前記スペクトルを徐々に変形させることを特徴とする音声信号変形接続方法。

【請求項６】　請求項５記載の音声信号変形接続方法において、

前記音声のスペクトルの変形に関しては、

前記２つの音声信号間の対応する音素内のピッチの対応を決定し、

対応したピッチ毎に、ある周波数を境界周波数として、一方の音声信号のスペクトルにおける前記境界周波数より高域の部分と、他方の音声信号のスペクトルにおける前記境界周波数より低域の部分とを結合させたスペクトルを生成し、生成されたスペクトルを当該ピッチにおけるスペクトルとすると共に、

前記境界周波数を単位時間毎に変化させることを特徴とする音声信号変形接続方法。

【請求項７】　請求項６記載の音声信号変形接続方法において、

前記境界周波数の変化は、単位時間毎に一定量増加するように行われることを特徴とする音声信号変形接続方法。

【請求項８】　請求項６記載の音声信号変形接続方法において、

2

前記境界周波数の変化は、変化開始時の値から変化終了時の値まで徐々に増加するように行われ、前記変化開始時に近い相対的に低域の段階では比較的ゆっくりと、前記変化終了時に近い相対的に高域の段階では比較的早く変化させることを特徴とする音声信号変形接続方法。

【請求項９】　請求項１記載の音声信号変形接続方法において、

前記パラメータは音声の基本周波数であり、前記所定の時間にわたって前記基本周波数を徐々に変化させることを特徴とする音声信号変形接続方法。

【請求項１０】　請求項９記載の音声信号変形接続方法において、

前記音声の基本周波数の変化に関しては、

前記各音声信号の平均基本周波数を求め、

両者の平均基本周波数の差とパラメータを変更する前記所定時間とに基づいて、単位時間当たりに変更すべき基本周波数の値を求め、

この値を変化量の単位として、一方の音声信号の平均基本周波数から他方の音声信号の平均基本周波数へと単位時間毎に変化させることを特徴とする音声信号変形接続方法。

【請求項１１】　請求項１記載の音声信号変形接続方法において、

前記パラメータは、音声のスペクトルと音声の基本周波数であり、

前記音声のスペクトルの変形に関しては、

前記２つの音声信号間の対応する音素内のピッチの対応を決定し、

対応したピッチ毎に、ある周波数を境界周波数として、一方の音声信号のスペクトルにおける前記境界周波数より高域の部分と、他方の音声信号のスペクトルにおける前記境界周波数より低域の部分とを結合させたスペクトルを生成し、生成されたスペクトルを当該ピッチにおけるスペクトルとすると共に、

前記境界周波数を単位時間毎に変化させるようにし、

前記音声の基本周波数の変化に関しては、

前記各音声信号の平均基本周波数を求め、

両者の平均基本周波数の差とパラメータを変更する前記所定時間とに基づいて、単位時間当たりに変更すべき基本周波数の値を求め、

この値を変化量の単位として、一方の音声信号の平均基本周波数から他方の音声信号の平均基本周波数へと単位時間毎に変化させるようにすることを特徴とする音声信号変形接続方法。

【請求項１２】　請求項１１記載の音声信号変形接続方法において、

前記音声のスペクトルと基本周波数とを並行して変化させることを特徴とする音声信号変形接続方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、録音編集型の音声メッセージの作成において、音声メッセージの追加、変更を効率よく行うことを可能とし、音声メッセージを用いたシステムの構築、維持の経済化をはかることができる音声信号変形接続方法に関する。

【０００２】

【従来の技術】現在、駅の構内アナウンスや、道路の渋滞などの情報を知らせるハイウェイラジオや、情報検索における音声ガイダンス等のサービスには、音声メッセージが使われている。これらの音声メッセージは、予め人間が発声した音声を録音し、この音声を継ぎはぎすることによって作成されている。

【０００３】係る音声メッセージの作成において、既に作成された音声メッセージとは異なる新たな音声メッセージが必要となり、その必要な音声メッセージが録音されていない場合には、新たに音声を追加録音する必要があった。この場合、既録音の音声と新規録音の音声との間で声質が急激に変化することなく自然につながるように、以前に発声した話者と同じ人に追加分を発声してもらう必要があった。

【０００４】

【発明が解決しようとする課題】しかしながら、同一話者であっても、以前の収録から年月が経っている等の理由で以前の声質と異なり、新旧メッセージの継ぎはぎにより聞き苦しさが予想される場合には、すべての音声メッセージを再び収録および作成し直す必要があった。また、以前に発声した人が不在の場合には、他の話者に代わりに発声してもらい、全ての音声メッセージを再び収録し直す必要があった。また、上記のような音声メッセージは音声合成装置を用いて作成することも可能であるが、この場合も、音声合成装置が異なる等の理由により互いに異なる声質となって出力された音声信号同士を接続する場合に、同様の問題が生じる。

【０００５】本発明は、このような背景の下になされたもので、声質の異なる音声メッセージ同士を違和感無く接続することができ、音声メッセージの追加、変更を効率よく行うことができる音声信号変形接続方法を提供することを目的とする。

【０００６】

【課題を解決するための手段】上記課題を解決するために、請求項１による音声信号変形接続方法においては、互いに異なる声質に属する２つの音声信号同士を接続する処理において、音声信号の特性を示すパラメータの値を、一方の音声信号の特徴を示す値から他方の音声信号の特徴を示す値へと所定の時間にわたって徐々に変化させることにより、２つの音声信号を接続することを特徴とする。

【０００７】また、請求項２による発明は、請求項１記載の音声信号変形接続方法において、前記パラメータの値を変更する所定時間にわたり、前記互いに異なる声質

の話者に同一のテキストを読み上げさせ、これにより得られる２種類の音声データを用いて前記パラメータ値の変更を行うことを特徴とする。

【０００８】また、請求項３による発明は、請求項１記載の音声信号変形接続方法において、前記互いに異なる声質の音声信号は音声合成装置による発声により得られるものであることを特徴とする。

【０００９】また、請求項４による発明は、請求項１記載の音声信号変形接続方法において、前記互いに異なる声質の音声信号は、一方は人間による発声、もう一方は音声合成装置による発声により得られるものであることを特徴とする。

【００１０】また、請求項５による発明は、請求項１記載の音声信号変形接続方法において、前記パラメータは音声のスペクトルであり、前記所定の時間にわたって前記スペクトルを徐々に変形させることを特徴とする。

【００１１】また、請求項６による発明は、請求項５記載の音声信号変形接続方法において、前記音声のスペクトルの変形に関しては、前記２つの音声信号間の対応する音素内のピッチの対応を決定し、対応したピッチ毎に、ある周波数を境界周波数として、一方の音声信号のスペクトルにおける前記境界周波数より高域の部分と、他方の音声信号のスペクトルにおける前記境界周波数より低域の部分とを結合させたスペクトルを生成し、生成されたスペクトルを当該ピッチにおけるスペクトルとすると共に、前記境界周波数を単位時間毎に変化させることを特徴とする。

【００１２】また、請求項７による発明は、請求項６記載の音声信号変形接続方法において、前記境界周波数の変化は、単位時間毎に一定量増加するように行われることを特徴とする。

【００１３】また、請求項８による発明は、請求項６記載の音声信号変形接続方法において、前記境界周波数の変化は、変化開始時の値から変化終了時の値まで徐々に増加するように行われ、前記変化開始時に近い相対的に低域の段階では比較的ゆっくりと、前記変化終了時に近い相対的に高域の段階では比較的早く変化させることを特徴とする。このような変化は人間の聴覚特性によりマッチしており、より自然な声質変化の実現を可能とする。

【００１４】また、請求項９による発明は、請求項１記載の音声信号変形接続方法において、前記パラメータは音声の基本周波数であり、前記所定の時間にわたって前記基本周波数を徐々に変化させることを特徴とする。

【００１５】また、請求項１０による発明は、請求項９記載の音声信号変形接続方法において、前記音声の基本周波数の変化に関しては、前記各音声信号の平均基本周波数を求め、両者の平均基本周波数の差とパラメータを変更する前記所定時間とに基づいて、単位時間当たりに変更すべき基本周波数の値を求め、この値を変化量の単

位として、一方の音声信号の平均基本周波数から他方の音声信号の平均基本周波数へと単位時間毎に変化させることを特徴とする。

【００１６】また、請求項１１による発明は、請求項１記載の音声信号変形接続方法において、前記パラメータは、音声のスペクトルと音声の基本周波数であり、前記音声のスペクトルの変形に関しては上記請求項６と同様の方法を用い、前記音声の基本周波数の変化に関しては上記請求項１０と同様の方法を用いることを特徴とする。

【００１７】また、請求項１２による発明は、請求項１１記載の音声信号変形接続方法において、前記音声のスペクトルと基本周波数とを並行して変化させることを特徴とする。

【００１８】

【発明の実施の形態】以下、図面を参照して、本発明の実施形態について説明する。図１（Ａ）～（Ｃ）は、互いに声質の異なる２人の話者による音声信号の波形と、これら音声信号を変形接続して得られた音声信号の波形との関係を示す波形図である。本実施形態の処理では、２人の話者（１及び２）に同じテキストを発声させ、これにより得られる音声信号（図１（Ａ）の１０１，図１（Ｂ）の１０２参照）を接続処理するものとする。

【００１９】処理の結果生成された音声信号は、図１（Ｃ）の符号１０３で示されるように、話者１の音声区間と変形接続区間と話者２の音声区間とから構成される。なお、この例では、２人の話者に同じテキストを発声させているが、必ずしも同じテキストを発声させる必要はない。例えば、以下に示す実施形態では、音声の特性を示すパラメータとして、基本周波数とスペクトルとを選び、該２つのパラメータにおける変形を行っているが、スペクトルの変形は行わず、基本周波数だけを変形する場合には、２人の話者が発声するテキストは異なっていても構わない。

【００２０】図２は、本実施形態による音声信号変形接続方法の全体の処理の流れを示すフローチャートである。２人の話者が発声した音声信号を入力すると、ステップＳ２０１では、それぞれの音声信号に音素境界を付与し、ステップＳ２０２に進む。ステップＳ２０２では、それぞれの音声信号に基本周期を示すピッチマークを付与し、ステップＳ２０３に進む。

【００２１】ステップＳ２０３では、上記ピッチマークに関して、両音声信号の対応した有声音区間に対し、最も近いピッチマーク同士を選ぶことにより、その区間のピッチマーク同士の対応を付ける。この結果、図３に破線で示すように、１対１、１対多、または、多対１の対応づけが得られる。このピッチマークの対応関係は、ピッチ対応テーブル３０１（図４参照）に記憶される。次に、ステップＳ２０４では、対応した音素ごとに音声信号パワーの正規化を行う。以上の処理は、音声収録後に独立して予め行ってもよいし、変形接続処理の一部として行ってもよい。

【００２２】次に、ステップＳ２０５では、後述する方法で音声信号の基本周波数の変換を行い、ステップＳ２０７に進む。ステップＳ２０７では、後述する方法で音声信号のスペクトルの変形を行い、ステップＳ２０９に進む。このとき、基本周波数の変更量の設定はステップＳ２０６で行われ、スペクトル変形に関与する境界周波数の変更量の設定はステップＳ２０８で行われる。これらの変更量は時間の関数となっている。

【００２３】最後に、ステップＳ２０９で、両音声信号が全体的に合成され、合成音を得る。上記ステップＳ２０５で使用可能な基本周波数の基本的変換方式としては、様々な方式が提案されているが、その一例としては、文献「E.Moulines, F.Charpentier, "Pitch-Synchronous Waveform Processing Techniques for Text-to-Speech Synthesis using Diphones", Speech Communication, Vol.9, pp.453-467, Dec.1990」で提案されているＰＳＯＬＡ方式がある。

【００２４】図４は、上記ステップＳ２０７で行われるスペクトル変形処理の一例を示すフローチャートである。同処理において、ステップＳ３０２では話者２の音声信号から、また、ステップＳ３０３では話者１の音声信号から、上記ステップＳ２０３（図２参照）で求められたピッチ対応テーブル３０１を参照して互いに対応するピッチを選択し、該ピッチ毎に、ピッチ同期信号に同期して音声波形を切り出す。ここでは、話者１から話者２へと徐々に音声を変形しながら接続する場合を例に挙げて説明する。この場合には、ピッチ同期の処理は、話者２のピッチマークの回数だけ行われる。ここで、図３の有声音Ｚの例に見られるように、話者１の２つのピッチマークが話者２の１つのピッチマークに対応している場合には、話者１の２つのピッチマークのうちどちらか一方を参照して音声波形を切り出す。一方、図３の有声音Ｙの例に見られるように、話者２の２つのピッチマークが話者１の１つのピッチマークに対応している場合には、話者１の１つのピッチマークにより参照される音声波形を２度切り出す。

【００２５】以下、１ピッチ分の処理を説明すると、ステップＳ３０４では、ステップＳ３０２で切り出した音声波形についてＦＦＴによるスペクトル分析を行う。また、ステップＳ３０４と並行して、ステップＳ３０５では、ステップＳ３０３で切り出した音声波形についてＦＦＴによるスペクトル分析を行う。ステップＳ３０６では、ステップＳ３０４で求めた話者２のスペクトルのうち、所定の周波数αＨｚより低い帯域の部分を取り出す。ステップＳ３０７では、ステップＳ３０５で求めた話者１のスペクトルのうち、上記周波数αＨｚより高い帯域の部分を取り出す。

【００２６】このステップＳ３０６およびＳ３０７で取

り出されたスペクトルは、ステップＳ３０８において、周波数αＨｚを境界にして結合される。このスペクトルの混合処理は、各ＦＦＴで得られたスペクトルの実部と虚部を、それぞれ個別に処理することで行われる。最後に、ステップＳ３０９では、両者のスペクトルを混合したスペクトルに対し、ＩＦＦＴを行い、１ピッチ波形を得る。こうして得られた１ピッチ波形は、上記ステップＳ２０９（図２参照）の処理に渡される。

【００２７】更に、上記境界周波数を時間的に変化させながら、ピッチ毎に同様のスペクトル混合処理を行うことにより、複数の”１ピッチ波形”が同様にステップＳ２０９の処理に渡され、最終的に該ステップＳ２０９で音声合成処理される。図５（Ａ）～（Ｃ）、図６（Ａ）～（Ｃ）、及び図７（Ａ）～（Ｃ）は、境界周波数の時間的変化に伴うスペクトル混合の例を示したものである。ここでは、境界周波数αを図５（Ｂ）→図６（Ｂ）→図７（Ｂ）に示す順序で３段階に変化させたとして、低域側が抽出される話者２の各段階のスペクトルを図５（Ａ）、図６（Ａ）、図７（Ａ）に、高域側が抽出される話者１の各段階のスペクトルを図５（Ｃ）、図６（Ｃ）、図７（Ｃ）に、そして、各段階でのスペクトル結合により得られた混合スペクトルを図５（Ｂ）、図６（Ｂ）、図７（Ｂ）にそれぞれ示す。

【００２８】図８は、本実施形態による音声信号変形接続処理において、ステップＳ２０５の基本周波数変換処理で変換される平均基本周波数と、ステップＳ２０７のスペクトル変形処理において変換される境界周波数との時間変化を示すグラフである。本実施形態では、音声の基本周波数変化の制御としては、話者１と話者２による音声信号の平均基本周波数をそれぞれ求めておき、両平均基本周波数の差とパラメータを変更する所定の時間（変形接続区間）とに基づいて単位時間当たりに変更すべき基本周波数値を求める。そして、該変形接続区間における基本周波数を、図８に示すように、上記２つの平均基本周波数の一方から他方へと、時間的に一定の割合で変化させる。

【００２９】また、音声のスペクトルの制御は、境界周波数αＨｚを時間的に一定の割合で変化させる。ここで、平均基本周波数の変更量はステップＳ２０６で設定され、また、境界周波数の変更量は、ステップＳ２０８で設定される。

【００３０】次に、図９（Ａ）～（Ｃ）は、図１（Ａ）～（Ｃ）に対応した、各音声信号による声紋を示すスペクトログラムである。各スペクトログラムにおいて、横軸は時間（ｓｅｃ）、縦軸は周波数（Ｈｚ）、そして、時間と周波数との各交点における濃さ（本紙面上では明確に現れないが）が、その時間におけるスペクトルの強さを表す。また、図９（Ｃ）に示す合成音声における変形接続区間には、該区間における境界周波数の変化を、符号１０５で参照されるラインで示した。

【００３１】なお、上述のような音声パラメータを変化させる割合は一定である必要はなく、様々な変化パターンが考えられる。例えば、図１０に境界周波数αの変化パターンの別例を示す。この例では、境界周波数が低い段階ではゆっくりと変化させ、境界周波数が高くなるにつれて早く変化させている。人間の聴覚は高域に比べて低域の周波数に重みがかった特性を有するため、この例のような変化をさせると、人間が聞いた時により一定の変化割合で、すなわち、より自然に声質を変化させることができる。

【００３２】以上、この発明の実施形態を図面を参照して詳述してきたが、具体的な構成はこの実施形態に限られるものではなく、この発明の要旨を逸脱しない範囲の設計の変更等があってもこの発明に含まれる。たとえば、上述した一実施形態においては、初めに基本周波数を変化させ、その後に音声スペクトルを変化させているが、変化の順番はこの逆でも構わないし、また、分散処理等により、両者を同時に変化させてもよい。なお、変形区間が長い場合には、これらパラメータの変化を分けて行った方が、より滑らかな接続が行える。

【００３３】また、変形接続される２つの音声信号は、人間による発声により得られたもの以外に、音声合成装置による発声で得られたもの同士でもよく、また、人間による発声と音声合成装置による発声で得られた音声信号を接続することも可能である。

【００３４】なお、以上の説明において、異なる話者の音声（メッセージ）を接続しても聞き手に違和感を与えない点を強調してきたが、本発明は、声の変化を聞き手に全く意識させないだけでなく、違和感の無い声の変化を聞き手に”意識させる”という利用法もある。例えば、画像処理のモーフィングと呼ばれる処理では、男性の顔の静止画像と女性の顔の静止画像とを用いて、（時間とともに画像を変化させながら）男性の顔を徐々に女性の顔へと変化させることが可能である。このような画像処理技術と本発明による方法とを統合すれば、人間の顔が男性から女性に変化しながら、その声もいつのまにか男性から女性に変化しているというような、見る（聞く）者に不思議な感覚を与えるシミュレーションの実現が可能である。このような技術は、映画やマルチメディア作品等の製作分野において、新しい表現手段として利用できる。

【００３５】

【発明の効果】上述のように本発明によれば、音声の特徴量を時間と共に変化させることができる。その結果、話者が異なる２つの音声を接続する場合であっても、接続区間における急激な声質の変化を避けることができ、聞く者にとって違和感なく音声を接続することが可能となる。

【図面の簡単な説明】

【図１】　本発明の一実施形態において、互いに声質の

9

異なる２つの音声の波形と、これらの音声信号に変形接続処理を施して得られた音声波形との関係を示す波形図である。

【図２】　同実施形態における音声信号変形接続方法の全体の処理の一例を示すフローチャートである。

【図３】　同実施形態において、２つの音声信号間のピッチマークの対応づけを説明するための図である。

【図４】　同実施形態におけるスペクトル変換処理の一例を示すフローチャートである。

【図５】　同実施形態において、ある時間における境界周波数の設定と、該境界周波数における２つのスペクトルの結合を説明するための図である。

【図６】　更に進んだ時間における境界周波数の再設定と、該境界周波数における２つのスペクトルの結合を説明するための図である。

10

【図７】　更に進んだ時間における境界周波数の再設定と、該境界周波数における２つのスペクトルの結合を説明するための図である。

【図８】　同実施形態による音声信号変形接続処理において、平均基本周波数と境界周波数の時間変化を示すグラフである。

【図９】　図１に示す各音声信号による得られる声紋を示すスペクトログラムである。

【図１０】　境界周波数の別の時間変化例を示すグラフである。

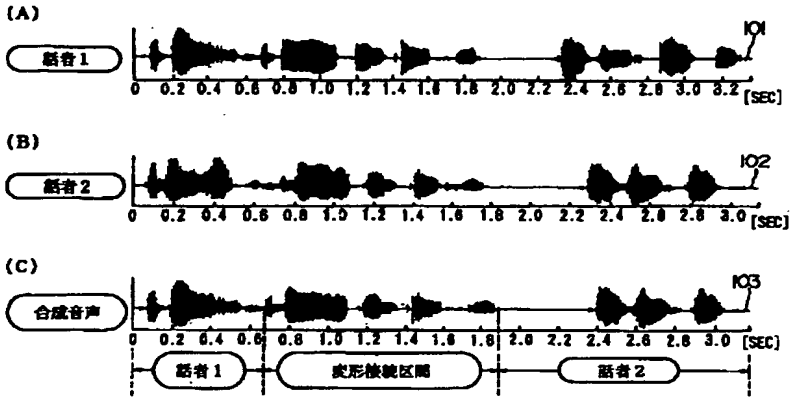
【符号の説明】

１０１　話者１の音声信号

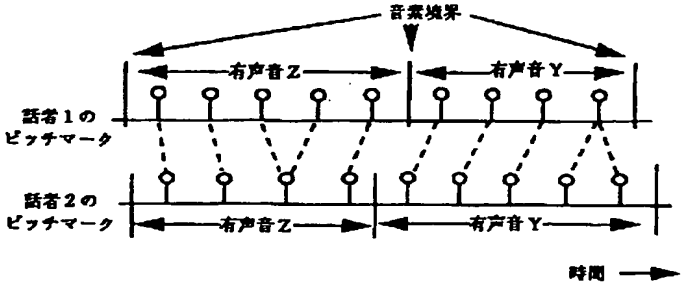
１０２　話者２の音声信号

１０３　合成音声信号

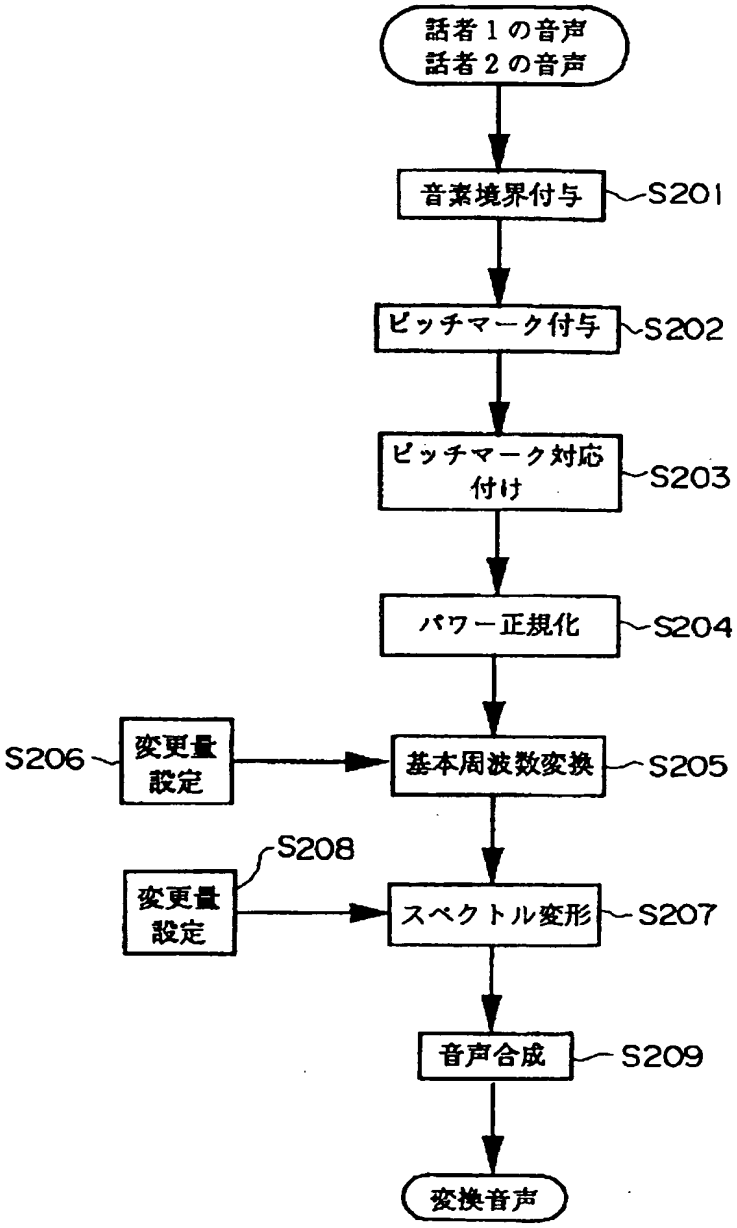
【図１】



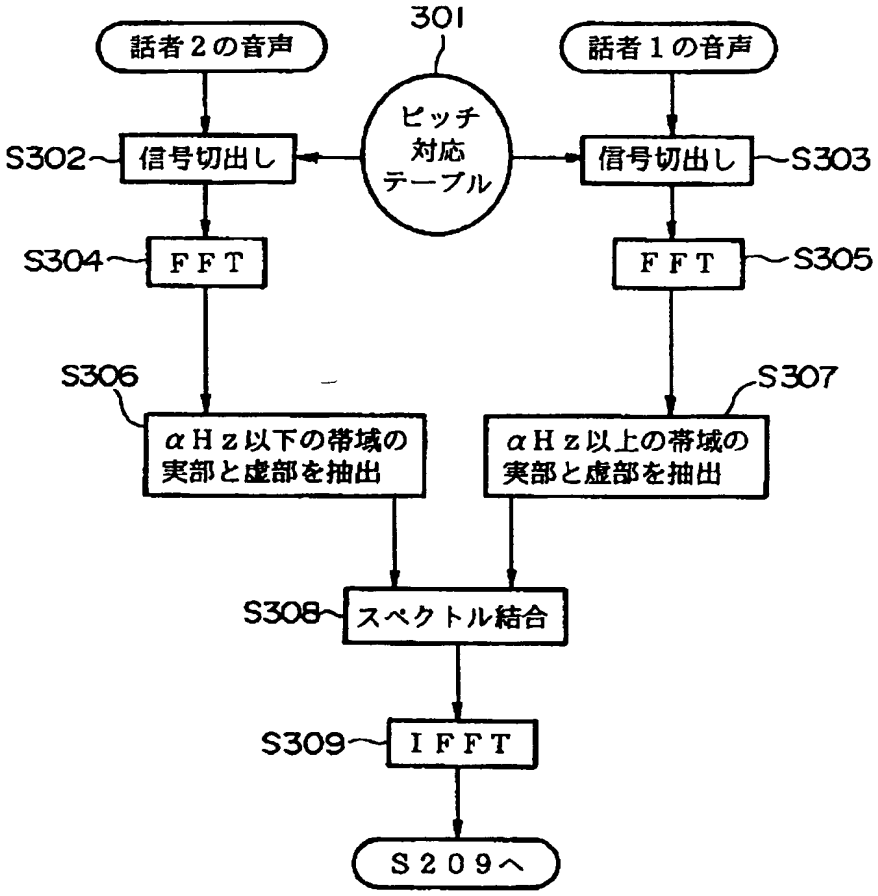
【図３】



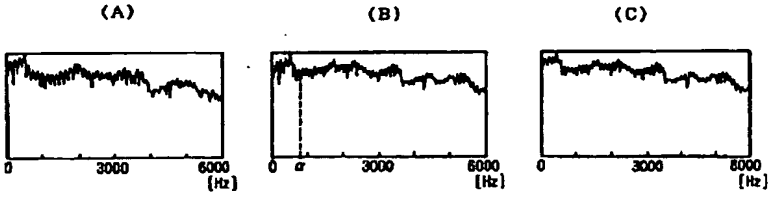
【図２】


話者１の音声
話者２の音声
音素境界付与
S201
ピッチマーク付与
S202
ピッチマーク対応付け
S203
パワー正規化
S204
S206
変更量設定
基本周波数変換
S205
S208
変更量設定
スペクトル変形
S207
音声合成
S209
変換音声

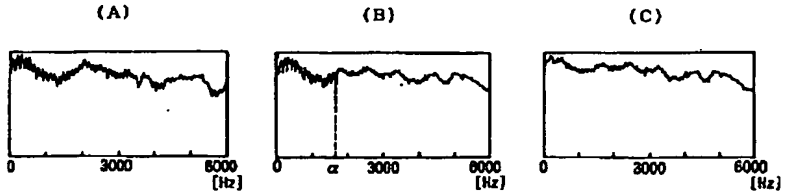
【図4】


301
話者２の音声
話者１の音声
ピッチ
対応
テーブル
S302
信号切出し
信号切出し
S303
S304
ＦＦＴ
ＦＦＴ
S305
S306
S307
αＨｚ以下の帯域の
実部と虚部を抽出
αＨｚ以上の帯域の
実部と虚部を抽出
S308
スペクトル結合
S309
ＩＦＦＴ
Ｓ２０９へ

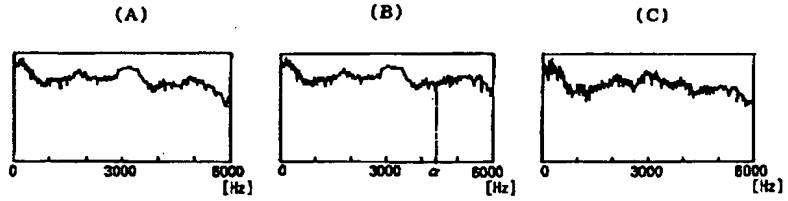
【図5】


(A)
(B)
(C)
0
3000
6000
[Hz]
α

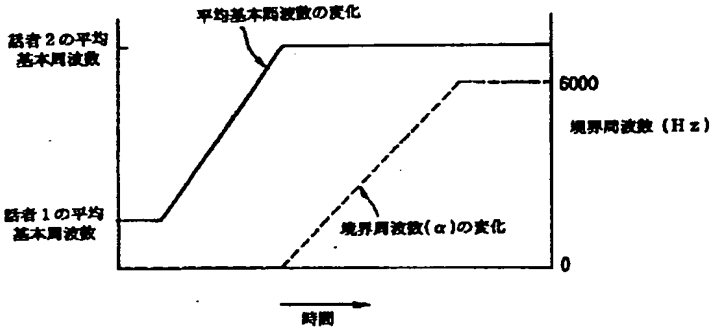
【図６】


(A)
(B)
(C)
0
3000
6000
[Hz]
α

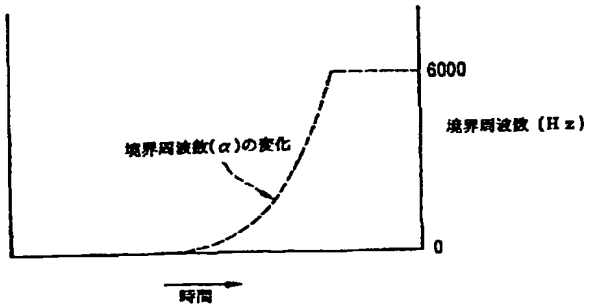
【図７】


(A)
(B)
(C)
0
3000
6000
[Hz]
α

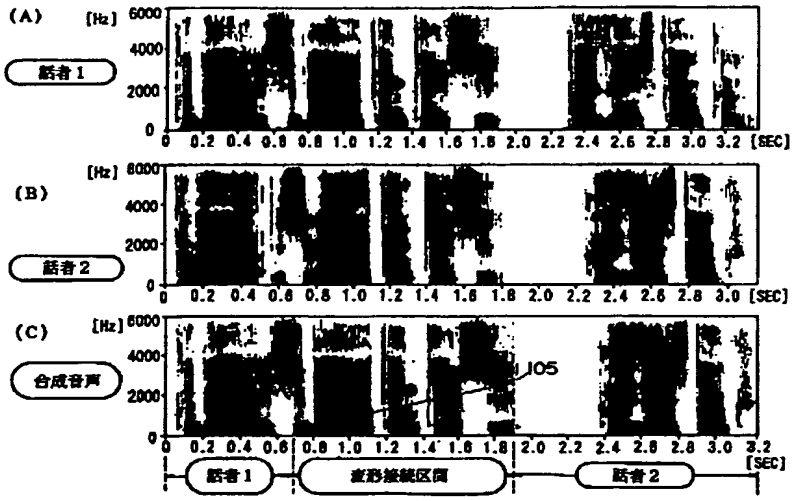
【図８】


平均基本周波数の変化
話者２の平均基本周波数
話者１の平均基本周波数
6000
境界周波数（Ｈｚ）
境界周波数(α)の変化
0
時間

【図１０】


6000
境界周波数（Ｈｚ）
境界周波数(α)の変化
0
時間

【図９】


(A)
[Hz]
6000
4000
2000
0
話者１
0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2 [SEC]
(B)
[Hz]
6000
4000
2000
0
話者２
0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 [SEC]
(C)
[Hz]
6000
4000
2000
0
合成音声
105
0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2
[SEC]
話者１
変形接続区間
話者２